"*In silico* Molecular Characterization of Begomovirus in Rajasthan, India"

A
SYNOPSIS
SUBMITTED

FOR THE FULFILMENT OF

Doctor of Philosophy (Ph. D.)

DEPARTMENT OF SCIENCE
FACULTY OF ARTS SCIENCE AND COMMERCE (FASC)
MODY INSTITUTE OF TECHNOLOGY AND SCIENCE (MITS)
(Deemed University u/s 3 of UGC Act, 1956)
LAKSHMANGARH, SIKAR (RAJASTHAN – 332311) INDIA

Under Supervision of:

Dr. Rajarshi Kumar Gaur
Ph.D (Biochemistry)
Dr. Vivek Bajpai
Ph.D (Microbiology)

Submitted by:

Rajneesh Prajapat
M.Sc. (Biotechnology), M.Phil

CONTENTS

S. NO	TITLE	PAGE NO.
1.	Origin of Proposal	
	• Introduction	1
2.	Review of Literature	5
3.	Methodology	10
4.	Significance of works	13
5.	Future prospects	13
6.	References	14

1. ORIGIN OF PROPOSAL

INTRODUCTION:

The term virus used by the ancient Romans to mean poison, venom or secretion – all unpleasant meaning. Viruses are the smallest biological structures that have the ability to replicate only inside living cells. Essentially they consist of a shell of protein which encloses a core of nucleic acid – either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA). There are over 700 known plant viruses, many of which cause catastrophic diseases and have wide host ranges. They have been classified into three families and 32 groups (Martelli, 1992). Serological and nucleic acid based techniques are being used for establish the identity or relatedness of plant viruses. Viruses with single stranded DNA that infect plants belong primarily to the family *Geminiviridae* and, a few of them, to the family *Circoviridae*.

Geminiviruses were recognized in 1978 by the International Committee on the Taxonomy of viruses on the basis of there unique virion morphology and possession of ssDNA as their genomic material (Matthews, 1979; Goodman, 1981). The family *Geminiviridae* is one of the largest plant virus family; its members have a circular, single-stranded DNA (ssDNA) genome of approximately 2.7–5.2 kb encapsulated within twinned (geminate) icosahedral virions. The protein coat of geminiviridae consists of one type protein molecule of about 28kd molecular weight. Based on their genome arrangement and biological properties, geminiviruses are classified into one of four genera: *Mastrevirus, Curtovirus, Topocuvirus* and *Begomovirus* (Stanley *et al.*, 2005). Whereas *Mastreviruses* and *Curtoviruses* are transmitted via various leafhopper , *Begomoviruses* are transmitted by the whitefly (*Bemisia tabaci*) and Topocuvirus is transmitted by the treehopper. The construction of the genome of geminiviruses, regardless of the genus to which they belong, is quite complex. Because geminiviruses replicate by producing temporary complementary (-) ssDNA strand, both the original virus (+) ssDNA and newly formed (-) ssDNA can have open reading frama (ORFs) that may be translated into fractional into fractional proteins (>10kd).Transcription of mRNA and subsequent translation into protein can take place in both directions in each of the ssDNAs.

Geminiviruses with two-part ssDNA genomes (Begomovirus) have a common 200 nucleotides sequence. They have open reading frames on the two (+) ssDNAs (DNA A and DNA B) and, in addition, on the two (-) ssDNAs. Circular open reading frames found in some geminiviruses correspond to identical or similar ORFs found in other geminiviruses, apparently because they

code for common functions in all viruses, e.g., replication, movement or vector transmissibility. However, some ORFs appear to be distinctive for the virus on which they are found.

Protein coded by the open reading frames include the coat protein, with a size of 27 to 31kd, that, in addition to provide the shell of the virus, is essential for infectivity in all single-component geminiviruses and play a role in insect vector specificity and transmissibility of the virus; the replication associated protein, which is usually the largest coded protein (40.3kd); and the protein facilitating cell-to-cell and systemic movement of geminiviruses seems to be coded by DNA-B of the virus.

Geminiviruses cause some of the most devastating disease of vegetables, such as tomato, bean, squash and field crops, like sugar beets, tobacco, corn and some medicinal important weeds. They are most common and catastrophic in the tropical and subtropical regions of the world because of the transmission by the Bemisia vector in these areas. Geminiviruses drastically reduce photosynthesis, plant growth, fruit growth and fruit quality losses, although they depend on the number of plats infected and on the age of plant at the time of infection, are frequency great. In recent years, the epidemics range observed from 30 to 100 % of the crops.

Begomoviruses are transmitted by the sweet potato whitefly, *Bemisia tabaci* (Gennadius). *B. tabaci* was first described in the genus Aleyrodes in 1889 (Gennadius, 1889), and was first reported as a pest in 1919 in India (Husain and Trehan, 1993). Since then, B. *tabaci* has been recognized as a pest of crops in tropical and subtropical countries. B. *tabaci* has a very wide host range, consisting of 500 species in 74 plant families (Greathead, 1986). The whitefly is a vector of viruses in the Geminiviridae, Potyviridae and Comoviridae families and the genera Carlavirus and Closterovirus.

Bemisia tabaci commonly found in tropical and sub-tropical regions, and in greenhouses of temperate areas. Rajasthan has Thar desert on its western part and weather has tropical or sub-tropical desert type of climate. Adult whiteflies look like tiny white moths, about 1-2 mm in length. They fly when the leaf is disturbed. The whitefly can acquire the virus after feeding on infected plants for 15 to 30 minutes, and can transmit the virus to another host plants after 24 hours of incubation within the insect. A period of at least 15 minutes feeding on the new host is subsequently required for transmission of the virus. Here the virus-vector relationship is mostly of circulative type (Costa, 1976).

Geminivirus infected plants often exhibit bright mottles to almost yellow mosaics and their leaves may be curled or otherwise distorted. When young plant becomes infected, they become dwarfed and infection of old plats result in reduced new growth, fruit set and new leaves appear

mottled. Control of geminiviruses is very difficult and primarily on measures that reduce the number of over seasoning insects and virus infected hosts; separating susceptible crop in place and time is recommended to reduce the transfer of viruliferous vectors from one crop to other. Few plant genes for only partial resistance to geminiviruses or to their vectors are available so far. Numerous crops are now genetically engineered with genes obtained from geminiviruses, with the hope of finding significant stable pathogen derived resistant to these viruses.

The largest genus, *Begomovirus*, currently contains 200 species (Fauquet *et al.*, 2008) that contains dicotyledonous infecting whitefly transmitted viruses in the family Geminiviridae, have either bipartite genomes (DNA-A and DNA-B) or monopartite genomes resembling DNA-A. DNA-A typically have six open reading frames (ORFs): AV1/V1 (coat protein, CP) and AV2/V2 (AV2/V2 protein) on the virion-sense strand, and AC1/C1 (replication initiation protein, Rep), AC2/C2 (transcriptional activator, TrAP), AC3/C3 (replication enhancer, REn) and AC4/C4 (AC4/C4 protein) on the complementary-sense strand. DNA-B has two ORFs, encoding movement proteins: BV1 (nuclear shuttle protein, NSP) on the virus-sense strand and BC1 (movement protein, MP) on the complementary-sense strand (Rojas *et al.*, 2005). Begomoviruses species those have bipartite genomes, comprising a DNA-A component required for replication and encapsidation and a DNA-B component required for virus movement (Hanley *et al.*, 2000). The International Committee on Taxonomy of Viruses database (ICTVdB) provide virus code to begomovirus, that is 00.029.0.03, virus accession number is 029003GE and NCBI Taxonomy ID is 10814 (Fauquet *et al.*, 2005).

The opposing transcription units of begomovirus DNA-A and DNA-B molecules are separated by an intergenic region (IR) that generally shares a highly conserved region of approximately 200 nt, named the common region (CR) (Lazarowitz, 1992). The CR contains an origin of replication (ori) that includes a stem–loop structure containing an invariant nonanucleotide (TAATATTAC) sequence whose T7–A8 site is required for cleaving and joining viral DNA during replication and conserved iterated sequences required for specific recognition and binding by Rep during replication (Arguello- Astorga *et al.*, 1994; Fontes *et al.*, 1994).

Based on phylogenetic studies and genome arrangement, begomoviruses have been divided broadly into two groups: the Old World (OW) viruses (Eastern hemisphere, Europe, Africa, Asia) and the New World (NW) viruses (western hemisphere, the Americas) (Padidam *et al.*, 1999; Paximadis *et al.*, 1999; Rybicki, 1994). Begomovirus genomes have a number of characteristics that distinguish Old World and New World viruses. All indigenous New World begomoviruses are bipartite, whereas both bipartite and monopartite begomoviruses are present

in the Old World. In addition, DNA-A of bipartite begomoviruses from the New World lacks an AV2 ORF (Rybicki, 1994; Stanley *et al.*, 2005). New World begomoviruses also have an N-terminal PWRSMAGT motif in the CP that is absent from Old World viruses (Harrison *et al.*, 2002). Until recently, it was thought that New World viruses arose more recently than Old World viruses, evolving after continental separation of the America from Gondwana (Rybicki, 1994). However, it was identified that virus indigenous to Vietnam, Corchorus yellow vein virus (CoYVV), resembles New World viruses more closely than Old World viruses (Ha *et al.*, 2008). This was based primarily on phylogenetic analysis, the absence of an AV2 ORF and the presence of an N-terminal PWRLMAGT motif in the CP.

Begomovirus nucleocaspid is 38 nm long and 15-22 nm in diameter. While particles have basic isocaherdal symmetry, they consist of two incomplete icosahedra. There are 22 capsomeres present per nucleocapsid and single stranded closed circular DNA also. The type species Potato yellow mosaic virus (PYMV) is a Begomovirus, first identified in the late 1980s. Disease is manifested in the infected plant as yellow mosaic or mottling, leaf distortion and crinkling and stunting. In this disease estimated yield loss of 50-60%. PYMV disease is also an economical problem in the Caribbean. The type species Bean golden yellow mosaic virus (BGYMV) also belongs to the Begomovirus genus. It causes a serious disease in bean species within Central America, the Caribbean and southern Florida.

PCR primers based on conserved sequences can be used to clone and sequence the N-terminal sequences of the CP and other genes of the geminiviruses; this sequence is sufficient to classify a virus isolate. A possible taxonomic structure for geminiviruses has been proposed after considering the sequence comparisons and biological properties (Malla *et al.*, 1995). Hot and dry conditions favour the whitefly and that help the spread of these viruses, therefore the conditions of Rajasthan (India) is suitable for begomoviruses transmission by whitefly because it is one of the driest and hottest regions in India. Whitefly populations decrease after heavy rains.

Considering all the facts mentioned above the following objectives taken for the present study:

1. To explore host range of Begomovirus in Rajasthan (India).
2. Amplification and sequencing of Begomovirus genes from host plant species using gene specific primers.
3. Comparative sequence analysis and phylogenetic study of isolated genes and sequences available in databases.
4. Comparative homology modeling and database designing.

2. REVIEW OF LITERATURE

Geminiviruses includes several early studies which had become classics in the literature of plant pathology. Curly top virus (CTV), the cause of major diseases in several crops in the western United States and the Mediterranean, has been the subject of considerable study since the early 1900s (Bennett, 1971).

The relationship of the leafhopper vector (*Circulifer tenellus*) to disease spread was understood before 1910, and significant progress was made before 1950 in developing control measures based upon genetic resistance in crop species, improvements in crop management and control of vectors with insecticides (Bennett, 1971; Goodman, 1981).

Geminiviruses were characterised by circular single stranded DNA (ssDNA) genomes encapsidated in twinned quasi isometric particles of about 18 x 30 nm. The *Geminiviridae* family divided into four genera: *Mastrevirus, Curtovirus, Topocuvirus* and *Begomovirus* (Van Regenmortel *et al.*, 1999).

Mastrevirus (type species *Maize streak virus*; acronym MSV) have monopartite genomes (2.6-2.8 kilobases) encoding 3 or possibly 4 proteins (ORF: V1, V2 and C1, C2). They were transmitted by leafhoppers (*Hemiptera: Cicadellidae*) in a persistent circulative, and non-propagative manner. They were most often found on monocotyledonous plants, especially on maize. Economically MSV causes the most devastating disease in maize in sub Saharan Africa resulting in up to 100% yield losses (Thottappilly, 1992). MSV was endemic in Africa where wild grasses are its natural hosts. The virus must moved to maize when the crop was introduced in Africa and neighbouring regions in the 16th century, as maize is also a congenial feeding host of *Cicadulina* spp., the vector of MSV (Varma and Malathi, 2003).

Curtovirus (type species *Beet curly top virus*; acronym BCTV) had monopartite genomes (2.9-3 kb) and phloem limited. The DNA codes for 6 or 7 proteins depending on the species. *Ludwigia hyssopifolia* showing yellow vein symptoms were collected from Nanning, Guangxi province of China. Viruses were detected by polymerase chain reaction (PCR) using universal primers for the genus *Begomovirus*. The complete DNA-A sequences of isolates obtained from each source plant comprised 2758 and 2745 nucleotides, respectively (Huang *et al.*, 2006).

The genus *Topocuvirus* created with a former curtovirus, *Tomato pseudo-curly top virus* as type species, on the ground of a different family of vector, treehopper (Hemiptera:

Micrutalis malleifera) instead of leafhopper , and on the fact that the sole representative had evolved by interspecies recombination between unknown viruses belonging to different genera (Briddon *et al.*, 1996). It was first reported in Florida (Stoner and Hogan, 1950). The TPCTV genome (2861 nts long) encodes 6 proteins (Briddon *et al.*, 1996) which were close to the curtoviruses.

Begomovirus (type species *Bean golden mosaic virus*) was the only genus of the family to be either bipartite with virus genes resident on two different circular ssDNA molecules (DNA-A, DNA-B) each of about 2.6-2.8 kb, or monopartite with all genes resident on one (DNA-A like) ssDNA of about 2.8 kb. The twinned particles had diameter of 18-20nm, 30 nm long, like most of the geminiviruses. The begomoviruses transmitted by the whitefly *B. tabaci* in a circulative manner and infecting dicotyledonous plants (Sunter and Bisaro, 1991).

Cotton leaf curl virus infection reduced the fecundity and longevity of *B. tabaci* compared with non-viruliferous whiteflies. Both viruliferous and non-viruliferous whiteflies survived longer on older plants than on younger plants. Female whiteflies survived longer than the male whiteflies irrespective of their being viruliferous or non-viruliferous. Plant age not affected the fecundity however egg viability declined with increasing age of plants (Sidhu *et al.*, 2009). Geminivirus DNA synthesis probably depends at least in part on enzymes provided by the host cell. DNA replication occurs by a mechanism based on a circular rather than a linear template (Dressier *et al.*, 1978).

The DNA A of the bipartite species is similar in arrangement to the genome of the monopartite begomoviruses. For New World bipartite begomoviruses, the DNA component lacks the V2 gene. The DNA-B component encodes BV1 and BC1, proteins that are essential for cell to cell and systemic movement (Noueiry *et al.*, 1994; Sanderfoot *et al.*, 1996), and can influence host range (Ingham *et al.*, 1995). Although not directly involved in interaction with the whitefly vector, DNA-B sequences affect the efficiency of virus acquisition by the insect by determining the location of begomoviruses in plant tissues (Liu *et al.*, 1997).

Mixed infection with two or more geminiviruses are common in tropical and subtropical areas (Torres-Pacheco *et al.*, 1996; Harrison, 1997; Sanz *et al.*, 2000; Pita *et al.*, 2001; Ribeiro *et al.*, 2003). Virus interactions during such mixed infections could play an important role leading to the development of complex diseases (Fondong *et al.*, 2000; Monci *et al.*, 2002). The type of interactions between co-infecting viruses can range from synergism where symptoms were stronger than those observed for the separate viruses (Kamei, 1969; Pruss *et al.*, 1997), to interference, where one virus affects the infection cycle of the second

virus (Malinshenko *et al.*, 1989; Fedorkin, 2000; Mendez Lozano *et al.*, 2003).

Possible interactions include also events such as complementation (Berrie *et al.*, 2001) or recombination (Sanz *et al.*, 2000; Pita *et al.*, 2001; Saunders *et al.*, 2001; Mendez Lozano *et al.*, 2003). An example of virus interactions as host dependent, and interference as antagonism was found between the *Pepper huasteco virus* and *Pepper golden mosaic virus*. Antagonism was found between the two viruses in pepper and in synergism when co-infected in *Nicotiana benthamiana* (Fondong *et al.*, 2000).

In addition to classical synergism between two geminiviruses new recombinants may arise, with more severe symptoms than induced by either parental virus. An example of this has occurred in Cameroon, where a new double recombinant had been formed between co-infecting *African cassava mosaic virus* isolates, developing more severe symptoms in the field compared to plants infected by either isolate alone (Fondong *et al.*, 2000). Another example of natural recombination of two begomoviruses observed with the *Tomato yellow leaf curl Sardinia virus* (TYLCSV) and TYLCV in Spain, the resulting recombinant appearing better fit than the two original isolates (Monci *et al.*, 2002).

The genetic diversity of begomovirus isolated from tomato (*Lycopersicon esculentum*) fields in the Southeastern region of Brazil was analyzed by direct sequencing of PCR fragments amplified by using universal oligonucleotides for the begomovirus DNA-A, and subsequent computer-aided phylogenetic analysis (Luciana *et al.*, 2002). Begomovirus-associated DNA-1 molecules are similar to nanovirus-encoded DNA-1 molecules in that they contain one major ORF that encodes a Rep protein and they replicate autonomously. The role of the begomovirus associated DNA-1, like molecules in disease progression is unclear (Briddon *et al.*, 2004).

The begomovirus associated diseases in cucurbits are emerging problems in India. For rapid and sensitive detection of begomovirus infection in cucurbits, nucleo-diagnostic methods had been developed. The nucleic acids hybridization test using radiolabelled probe to putative coat protein gene of *Tomato leaf curl New Delhi virus*-[Luffa] (ToLCNDV-[Luf]) detected begomovirus in several cucurbits and whitefly. The method was efficient in detecting begomovirus in crude extracts of field samples. Polymerase chain reaction (PCR) was standardised using primers synthesised from AC1 and AV2 (coat protein [CP]) genes of ToLCNDV-[Luf]. The CP gene based primers successfully detected begomovirus in naturally infected several cucurbits. Both the molecular tools were validated for broad spectrum diagnosis of begomovirus infection in eight cucurbit crops. Emergence of begomovirus infection in bottle gourd (*Lageneria siceraria*), ivy gourd (*Coccinia cordifolia*), pumpkin (*Cucurbita moschata*), ridge gourd (*Luffa acutangula*)

and watermelon (*Citrullus lanatus*) in northern India had been identified (Sohrab *et al.*, 2006).

A *Begomovirus* causing yellow vein mosaic disease of pumpkin (*Cucurbita maxima* L.) was characterized at molecular level by cloning and sequence analysis of its complete DNA-A genome. The DNA-A of the isolate contains 2758 nucleotides which encode six open reading frames (ORFs): AV1 and AV2 in the virion-sense and AC1, AC2, AC3 and AC4 in the complementary-sense. Based on the highest (96%) sequence identities and close phylogenetic relationships with *Squash leaf curl China virus* species, the *Begomovirus* was identified as strain of *Squash leaf curl China virus*. The presence of DNA-B genome of the virus strain was also detected by dot blot hybridization test using DNA-B specific probe (Singh, 2008).

Association of *Begomovirus* with the disease could be confirmed by sequence analysis. Amplification of the virus was conducted following whitefly (*Bemisia tabaci* Genn.) transmission. High degree of genetic diversity among Begomovirus group, which might be emerged due to a high degree of preexisting genetic diversity among begomoviruses or possible recombination between different virus species during infection (Ribeiro *et al.*, 2003; Sri *et al.*, 2008).

The soil-borne plant pathogen *Agrobacterium tumefaciens* supported the replication of tomato leaf curl geminivirus (Australian isolate) (TLCV) DNA. Replicative DNA forms of three other geminiviruses also accumulated at varying levels in *Agrobacterium*. Geminiviral DNA constructs that lacked the ability to replicate in *Agrobacterium* were rendered replication-competent by changing their configuration so that two copies of the viral *ori* were present. The low-level replication of TLCV DNA can occur in *Escherichia coli* containing a dimeric TLCV construct in a high copy number plasmid (Selth *et al.*, 2002).*Ageratum conyzoides* exhibiting yellow vein symptoms, collected near Haikou, Hainan Province, China, contained begomoviral DNA-A like molecules. There is less specific interaction between monopartite begomovirus and DNAβ than between the DNA-A and DNA-B of begomoviruses with bipartite genomes (Xiong *et al.*, 2006).

Yellow vein mosaic disease had the major limitation in the production of bhendi or okra (*Abelmoschus esculentus*), an important vegetable crop of India. This disease caused by a complex consisting of the monopartite begomovirus *Bhendi yellow vein mosaic virus* (BYVMV, family: *Geminiviridae*) and a small satellite DNA β component. BYVMV can systemically infect bhendi upon agroinoculation but produces only mild leaf curling in this host. DNA β induces typical symptoms of bhendi yellow vein mosaic disease (BYVMD) when co-agroinoculated with the begomovirus to bhendi. The DNA β component associated with BYVMD had a number of

features in common with those reported for ageratum yellow vein disease and cotton leaf curl disease. BYVMV represents a new member of the emerging group of monopartite begomoviruses requiring a satellite component for symptom induction (Jose and Usha, 2003).

Cloned DNA-B components, belonging to the bipartite begomoviruses *Indian cassava mosaic virus* (ICMV) and *Sri Lankan cassava mosaic virus* (SLCMV), family *Geminiviridae*, when co-inoculated along with previously cloned DNA-A components of the respective viruses onto the experimental host *Nicotiana benthamiana*, generated defective DNAs (def-DNA) ranging in size from 549 to 1555 nucleotides. All the cloned def-DNAs contained the common region (CR) as well as portions of either DNA-A or DNA-B and, in a few cases, both DNA-A and DNA-B, representing recombinant products, the junction points of which correspond to repeats of 2–11 bases found in the parental molecules. The DNA-B-derived def-DNAs were, in some cases, associated with a decrease in levels of DNA-B, with a concomitant change in the symptoms from downward leaf curling in the older leaves to upward leaf-rolling in newly emerging leaves, more typical of monopartite begomoviruses (Patil *et al.*, 2007).

Approximately 60 different geminiviruses had been reported to be transmitted by B. *tabaci* (Markam *et al.*, 1994). In the New World before 1986, B. *tabaci* was considered a pest of a limited number of crops (tobacco, cotton, potato, bean, soybean), but by 1986 a sudden increase of the whitefly population in ornamentals in Florida was observed (Osborne, 1988). Shortly after that, whiteflies were reported in other crops in Florida, in California (Perring *et al.*, 1991), and in Texas, Arizona, Central America and South America (Brown, 1994). In Florida the whitefly infestation was associated with silver leaf of squash and irregular ripening of tomato (Maynard and Cantliffe, 1989).

Yellow vein mosaic disease of mesta (*Hibiscus* spp.) poses a serious threat to the cultivation of this crop in India. The disease was found to be associated with two different whitefly-transmitted monopartite begomoviruses, *Mesta yellow vein mosaic virus* and *Mesta yellow vein mosaic Bahraich virus*, together with two betasatellite species, *Cotton leaf curl Multan betasatellite* and *Ludwigia leaf distortion betasatellite* (Roy *et al.*, 2009). New virus species names assigned by using the rules proposed by the ICTV *Geminiviridae* Study Group. Demarcation of new viral species based on DNA-A sequence comparisons, using a threshold of 89% nucleotide sequence identity (Fauquet *et al.*, 2008). The sequence and arrangement of introns, Rep N-terminal iterated-related domains (IRD) (Arguello-Astorga *et al.*, 1994) and recombinant regions were also considered in the taxonomic decisions. A sanctioned taxonomic system had not available for naming the begomovirus associated satellite molecules (DNA-b and DNA-1) (Zhou *et al.*, 2003).

3. METHODOLOGY

(i) Collection of plant samples: The leaf samples will be collected from the different regions of Rajasthan (India) on the basis of typical begomovirus symptoms i.e., vein yellowing, leaf curling, chlorosis and stunting.

(ii) DNA extraction and PCR amplification: Total DNA from leaf samples will be extracted by using a CTAB (Cetyl Trimethyl Ammonium Bromide) method. The plant material will be suspended in 900 µl of 3×CTAB buffer (3% cetyl trimethyl ammonium bromide [CTAB]; 1.4 M NaCl; 100 mM Tris-HCl, pH 8.0; 20 mM EDTA and 0.2% β-mercaptoethanol), then will be incubated for 30 min at 65°C in waterbath. The lysates will be extracted twice with equal volume of chlorophorm: isoamyl alcohol (24:1). DNA will be precipitated with 1 volume of 100% isopropanol (30 min at -20°C, followed by centrifugation at 12000 rpm for 10 min). The DNA pellet will be washed with cold 70% ethanol twice, vacuum dried, and resuspended in 100 µl of TE buffer (Manen *et al.*, 2005).Concentration of DNA will be estimated through NanoDrop (Cuvette free spectrophotometer) to check the quantity of DNA per microliter for further study.

PCR will be carried out in 50 µl volumes by using begomovirus specific primer (CP). The presence/absence of a DNA-A genome component will also determine for each of the isolates using the begomovirus specific PCR primers. This will be confirming the molecular level infection of the begomovirus in plant samples.

After PCR amplification the gel electrophoresis technique will be used for the separation of deoxyribonucleic acid (DNA). 1 % agarose will be prepared in 1 ml TAE buffer with 2 µl Ethidium Bromide (EtBr) for making electrophoresis gel.

(iii) Sequencing of DNA samples: Sequencing of the amplified product will be carried out by using the M13 universal primer. In order to obtain larger amounts of DNA for the sequencing reaction, new PCR reactions will be carried out in a 100 µl volume, with the same components described above and using the standard amplification profile.

(iv)Submission of sequences in National Centre for Biotechnology Information (NCBI) and Homology search: After sequencing the DNA sequence will be submitted to NCBI-BankIt or Sequin. Homology search will be carryout by using alignment tools (e.g., NCBI-BLAST etc.).

(v) Phylogenetic study of sequences by using Bioinformatics tools: Pair wise sequence analysis will be carried out by using the GCG software package (Devereux *et al.*, 1984) or other alignment tools. Database search will be carried out with BLAST*n* or with other BLAST tools (http://www.ncbi. nlm.nih.gov/BLAST). Multiple sequence alignments will be prepared with CLUSTAL-W (http://www.ebi.ac.uk/clustalw).

Sequences will be aligned with the CLUSTAL-X program (Thompson *et al.*, 1997) or Clustal-W (www.ebi.ac.uk/clustalw) and phylogenetic trees will be constructed by using the Neighbour-Joining algorithm and viewed with the TreeView program (Page, 1996) or with MEGA version 2.0 (http://www.megasoftware.net).

(vi) Homology Modeling for three dimensional (3D) structure construction and stability study of proteins:

- There are several DNA and protein sequences of begomovirus available at Gen-Bank NCBI, ExPASy and other database those 3D structures have been unknown, therefore by using homology modelling principles, the 3D model of various begomovirus proteins will be constructed (modeling will be carry out only for the sequences those show high degree of identity/similarity during alignment, e.g., BLAST). To determine this, various server and tools will be utilized like ExPASy (Expert Protein Analysis System), NCBI-BLASTp , PDB (Protein Data Bank), 3D-JIGSAWN (Protein Comparative Modeling Server), Verify 3D Structure Evaluation server, UCLA-DOE server, SAVA (Structure Analysis and Validation server) and Chimera (Gaur *et al.*, 2009) according to the protocol steps. Results will be used for comparative energy stability and molecular phylogenetic study. Final results will be further analysed with respect to our sequencing results and for secondary database designing (Kasturi and Kumar, 2010; Prajapat *et al.*, 2010).
- DNA sequencing results will be further utilized for homology modeling procedure according to previously explained protocol by using on-line and off-line bioinformatics tools. Final results will be further used for primary database designing.

(vii) *In-Silico* position specific mutation study: Proteins and nucleotides sequence information will be further analysed to find out site specific mutation, conserved gene mutation (e.g., position of tyrosine etc.) by using bioinformatics tools (e.g., Chimera or other software).

(viii) **Database designing (Primary and Secondary):** Resulting data will be used for the database designing on MySQL by using PHP (Hypertext Preprocessor) as front end and the computing platform on Linux by using Apache web server as middle tire.

(ix) **Statistical analysis:** Analysis of data will be done either by using Bivariate or Multivariate statistics techniques or the choice of techniques will be depend on availability of resulting data.

(x) Publication of research results and thesis writing.

4. SIGNIFIANCE OF WORK

(a) Hot and dry climatic conditions of Rajasthan favour the high population of whitefly (*Bemisia tabaci*). This study gives the information about the natural host plants and transmission vector in Rajasthan (India).

(b) Several weeds in Rajasthan serve as the natural host of Begomovirus. Weeds show biological competition with the field crops and in several cases allows the transmission of Begomovirus from weeds to the field crops. This will increase host range of Begomovirus and yield loss from 80 to 100 %.

(c) Begomovirus genomics and proteomics database helpful to the understand the nature of virus and genetic diversity in Rajasthan region.

(d) Relationship between Begomovirus infection and the yield loss of economically important crops, medicinal weeds and other host plant species.

(e) Information about role of Begomovirus genes in disease symptoms and disease development.

(f) Information about phylogenetic and evolutionary relationship of Indian Begomovirus strains with other world wide distributed stains.

5. FUTURE PROSPECTUS

(a) Results can be used to develop resistance tansgenic plants using RNAi.

(b) Basic information about making quarantine strategies at state and national level for Begomovirus information control.

(c) Role of particular single gene in disease development mechanism can be understand by using genomics and proteomics data.

(d) Genetic diversity information among the members of geminivirus, infecting different host plants in Rajasthan will be considered for designing disease control strategies by using chemical agents.

(e) Begomovirus database information can be helpful in the research of *in-silico* ligand molecules designing for virus infection control.

6. REFERENCES

1. Arguello-Astorga, G.R., Guevara-Gonzalrz, R.G., Herrera-Estrella, L.R., and Rivera-Bustamante, R.F. (1994). Gemenivirus replication origine have a group specific organization of iterative elements: A model for replication. *Virology*, **203**: 302-1000.
2. Berrie, L.C., Rybicki, E. P. and Rey, M. E. C. (2001). Complete nucleotide sequence and host range of South African cassava mosaic virus: further evidence for recombinations amongst begomoviruses. *Journal of General Virology*, **82**: 53-58.
3. Bennett, C. W. (1971). The Curly Top Disease of Sugarbeet and Other Plants, Monograph No. 7. St. Paul, Minnesota: American Phytopathological Society.
4. Brown, J. K. (1994). Current status of *Bemisia tabaci* as a plant and virus vector in agroecosystems worldwide. *FAO Plant Prot. Bull*, **41**: 3-32.
5. Briddon, R.W., Bull S.E., Mansoor S., Amin I., Mansoor, S., Bedford I.D., Rishi, N., Siwatch, S.S., Zafar Y., Salam, A.M.A. and Markham, P.G. (2004).Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus–DNA β complexes. *Virology*, **324**(2): 462-474
6. Briddon, R. W., Bedford, I. D., Tsai, J. H. and Markham, P. G. (1996). Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. *Virology*, **219**: 387-394.
7. Brunt, A.A., Crabtree, K., Dallwitz, M.J., Gibbs, A.J. and Watson, L. (1996). Viruses of Plants: Descriptions and lists from the VIDE database. CAB International, Cambridge, UK, pp. 1484.
8. Costa, A.S. (1976). Whitefly transmitted plant diseases. *Ann. Rev. Phytopath*, **14**: 429-449.
9. Devereux, J., Haeberli, P. and Smithies, O.A. (1984). Comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research*, **12**:387-395.
10. Dressler, D., Hourcade, D., Koths, K. and Sims, J. (1978). The DNA replication cycle of the isometric phages. In The Single Stranded DNA Phages. New York: Cold Spring Harbor Laboratory, pp. 187-214.
11. Fauquet, C.M., Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M. and Zhou, X. (2008). Geminivirus strain demarcation and nomenclature. *Archives of Virology*, **153**: 783–821.
12. Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U. and Ball, L.A. (2005). Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, pp. 1259.
13. Fedorkin, O.N., Merits, A., Luccesi, J., Solovyev, A.G., Saarma, M., Morozov, S.Y. and Makinen, K. (2000). Complementation of movement deficient mutations in popato virus X: Potyvirus coat proteins mediates cell to cell traficking of C-terminal truncation but not deletion mutant of potexvirus at protein. *Virology*, **270**: 31-42.
14. Fondong, V.N., Pita, J.S., Rey, M.E.C., DeKochko, A., Beachy, R.N. and Fauquet, C.M. (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *Journal of General Virology*, **81**: 287-297.
15. Fontes, E.P., Eagle, P.A., Sipe, P.S., Luckow, V.A. and Hanley-Bowdoin, L. (1994). Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *J Biol Chem*, **269**: 8459–8465.
16. Funayama, S. (2001). Effects of Virus Infection and Light Environment on Population Dynamics of Eupatorium makinoi (Asteraceae). *American Journal of Botany*, **88**: 616.
17. Gaur, R.K. and Prajapat, R. (2009). Approaching online tools of Bioinformatics to identify the 3D structure of Plastocyanine protein of *Ginkgo biloba* (Ginkgo) using Homology Modeling. *IJBER*, **2**:77-86.
18. Gennadius, P. (1889). Disease of tobacco plantations in the Trikonia. The aleurodid of tobacco. *Ellenike Georgia*, **5**:1-3.
19. Goodman, R.M. (1981). Geminiviruses. *J.gen. Virol*, **54**: 9-21.
20. Goodman, R. M. (1981). Geminiviruses: the single-stranded DNA viruses of plants. In Handbook of Plant Virus Infections. Amsterdam: North-Holland/Elsevier.

21. Greathead, A.H. (1986). Host plants. In Bemisia tabaci-A literature Survey. M. J. W. Cock (Ed.), CAB International Institute of Biological Control, Silwood Park, Ascot, Berks., UK, pp.17 - 25.
22. Ha, C., Coombs, Steven, R.P., Harding, R., Vu, M. and James, D. (2008.) Molecular characterization of begomoviruses and DNA satellites from Vietnam: additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. *Journal of General Virology*, **89**: 312–326.
23. Hanley, B.L., Settlage, S.B., Orozco, B. M., Nagar, S. and Robertson, D. (2000). Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Biochemistry and Molecular Biology*, **35**: 105-140.
24. Harrison, B.D., Swanson, M.M. and Fargette, D. (2002). Begomovirus coat protein: serology, variation and functions. *Physiological and Molecular Plant Pathology*, **60** (5): 257-271.
25. Harrison, B. D., Zhou, X., Otim-Nape, G.W., Liu, Y. and Robinson, D.J. (1997). Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. *Annals of Applied Biology*, **131**: 437-448.
26. Huang, J.F., Jiang, T. and Zhou, X.P. (2006). Molecular characterization of Begomovirus infecting Ludwigia hyssopifolia. *Journal of Plant pathology*, **88** (1): 83-88.
27. Husain, M.A. and Trehan, K.N. (1993). Observations on the life-history, bionomics and control of white-fly of cotton (Bemisia gossypiperda M. & L.). *Indian Journal of Agricultural Science*, **3**:701 - 753.
28. Ingham, D.J., Pascal, E. and Lazarowitz, S. G. (1995). Both bipartite geminivirus movement proteins define viral host range, but only BL1 determines viral pathogenicity. *Virology*, **207**: 191-204.
29. Jose, J. and Usha, R. (2003). Bhendi Yellow Vein Mosaic Disease in India Is Caused by Association of a DNA β Satellite with a Begomovirus. *Virology*, **305** (2): 310-317.
30. Kamei, T., Goto, T. and Matsui, R. (1969). Reduced turnip mosaic virus multiplication in leaves infected with cauliflower mosaic virus. *Phytopathology*, **59**: 1513-1516.
31. Kasturi, K., Kumar, S.C. (2010). Homology modelling and structure based validation of *Mtd-Ddla* of *M. tuberculosis*. *Res. J. Biotech*, **5** (2): 43-50.
32. Liu, S., Bedford, I. D., Briddon, R. W. and Markham, P. G. (1997). Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. *Journal of General Virology*, **78**: 1791-1794.
33. Lazarowitz, S.G. (1992). Geminiviruses: genome structure and gene function. *Crit Rev Plant Sci*, **11**: 327–349.
34. Luciana, P.A., Renata F. C., Elizabeth, P.B.F., Murilo, G.D.C. and Murilo, Z.F. (2002). Genetic diversity of Begomovirus Infecting Tomato and associated weeds in Southeastern Brazil. *Fifopatol. bras*, **27** (4): 372-377.
35. Malla, P., Roger, N.B. and Claude, M.F. (1995). Classification and identification of geminiviruses using sequence comparisons. *J Gen Virol*, **76**: 249-263.
36. Malinshenko, S. I., Kondakova, O. A., Taliansky, M. E. and Atabekov, J. G. (1989). Plant virus transport function: complementation by helper viruses is non-specific. *Journal of General Virology*, **70**: 2751-2757.
37. Manen, J. F., Olga, S., Lorène, A., Alexander, V. M. and Arkady, S. (2005). A fully automatable enzymatic method for DNA extraction from plant tissues.*BMC Plant Biology*, **5**:23.
38. Matthews, R.E.F. (1979). Classification and nomenclature of viruses. *Intervirology*, **12**: 129-296.
39. Markam, P. G., Bedford, S. L., and Pinner, M.S. (1994). The transmission of geminivirus by Bemisia tabaci. *Pestic. Sci*, **42**: 123 - 128.
40. Martelli, G. P. (1992). Classification and nomenclature of plant viruses – state of the art. *Plant Disease*, **76**: 436– 442.
41. Martin, D. and Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics*, **16**: 562–563.
42. Martin, D. P., Willment, J. A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., James, D. and Rybicki, E. P. (2001). Sequence diversity and virulence in *Zea mays* of *Maize streak virus* isolates. *Virology*, **288**: 247- 255.

43. Maynard, D.N. and Cantliffe D.J. (1989). Squash silverleaf and tomato irregular ripening: new vegetable disorders in Florida. University of Florida, IFAS, Florida Cooperative Extension Service, Vegetable Crops Fact Sheet VC-37.
44. Mendez L. J., Torres P. I., Fauquet, C. M. and Rivera Bustamante, R. F. (2003). Interactions between geminiviruses in a naturally occurring mixture: Pepper huasteco virus and Pepper golden mosaic virus. *Phytopathology*, **93**:270-277.
45. Monci, F., Sanchez-Campos, S., Navas-Castillo, J. and Moriones, E. (2002). A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. *Phytopathology*, **303**: 317-326.
46. Noueiry, A.O., Lucas, W.J. and Gilbertson, R.L. (1994). Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell*, **76**: 925-932.
47. Osborne, L.S. (1988). The not so sweet sweetpotato whitefly. *Florida Foliage*, **14** (5):8-15.
48. Page, R.D.M. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*, **12**: 357–358.
49. Perring, T. M., Cooper, A.D., Rodriguez, R. J., Farrar, C. A. and Bellows, T.S. (1991). Identification of a whitefly species by genomic and behavioral studies. *Science*, **259**: 74-77.
50. Patil B. L., Nitin, D., Rob, W. B., Simon, E. B., Dirk, R., Basanta, K. B., Indranil, D., John S. and Holger, J. (2007).Deletion and recombination events between the DNA-A and DNA-B components of Indian cassava-infecting geminiviruses generate defective molecules in *Nicotiana benthamiana*. *Virus Research*, **124** (1-2):59-67.
51. Pita, J. S., Fondong, V. N., Sangare, A., Kokora, R. N. N. and Fauquet, C. M. (2001). Genomic and biological diversity of the African cassava geminiviruses. *Euphytica*, **120**: 115-125.
52. Paximadis, M., Idris, A.M., Torres-Jerez, I., Villarreal, A., Rey, M.E.C. and Brown, J.K. (1999). Characterization of tobacco geminiviruses in the Old and New World. *Arch Virol*, **144**: 703–717.
53. Prajapat, R., Gaur, R.K., Raizada, R. and Gupta, V.K. (2010). *In Silico* analysis of Genetic diversity of Begomovirus using Homology Modelling. *Journal of Biological Science*, **10** (3): 217-223.
54. Pruss, G., Ge, X., Shi, X. M., Carrington, J. C. and Vance, V. B. (1997). Plant viral synergism: The potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell*, **9**: 859-868.
55. Ribeiro, S. G., Ambrozevicius, L. P., Avila, A. C., Bezerra, I.C., Calegario, R. F., Fernandes, J. J., Lima, M. F., Mello, R. N. d., Rocha, H. and Zerbini, F. M. (2003). Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. *Archives of Virology*, **148**: 281-295.
56. Rybicki, E.P. (1994). A phylogenetic and evolutionary justification for 3 genera of Gemeniviridae. *Arch. Virol*, **139**: 49-77.
57. Rojas, M.R., Hagen, W.J.L. and Gilbertson, R.L. (2005). Exploiting chicks in the plant armor. Evolution and emergence of gemeniviruses. *Annu. Rev. Phytopathol*, **43**: 361-394.
58. Roy, A., Acharyya, S., Das, S., Ghosh, R., Paul, S., Srivastava, R. K. and Ghosh, S. K. (2009).Distribution, epidemiology and molecular variability of the begomovirus complexes associated with yellow vein mosaic disease of mesta in India. *Virus Research*, **141** (2) :237-246
59. Sanderfoot, A.A., Ingham, D.J. and Lazarowitz, S. G. (1996). A viral movement protein as a nuclear shuttle – The geminivirus BR1 movement protein contains domains essential for interaction with BL1 and nuclear localization. *Plant Physiol*, **110**: 23-33.
60. Sanz, A.I., Fraile, A., GarciaArenal, F., Zhou, X.P., Robinson, D.J., Khalid, S., Butt, T. and Harrison, B.D. (2000). Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. *Journal of General Virology*, **81**: 1839-1849.
61. Saunders, K., Bedford, I.D. and Stanley, J. (2001). Pathogenicity of a natural recombinant associated with Ageratum yellow vein disease: Implications for geminivirus evolution and disease aetiology. *Virology*, **282**: 38-47.
62. Selth, L.A., John, W.R. and Rezaian, M.A. (2002). *Agrobacterium tumefaciens* supports DNA replication of diverse geminivirus types. *FEBS Letters*, **516** (1-3):179-182.

63. Sidhu, J.S., Mann, R.S. and Butter, N.S. (2009). Deleterious effects of cotton leaf curl virus on longevity and fecundity of whitefly, *Bemisia tabaci* (Gennadius). *J. Entomol*, **6**: 62-66.

64. Singh, R., Raj, S. K. and Prasad, V. (2008).Molecular Characterization of a Strain of *Squash leaf curl China Virus* from North India. *Journal of Phytopathology*, **156** (4): 222-228.

65. Sri, H.H., Orawan, C. and Noor, A. (2008). Molecular Identification and sequence analysis of *Tobacco Leaf Curl Begomovirus* from Jember, East Java, Indonesia. *Hayati Journal of Biosciences*, **15** (1): 13-17.

66. Stanley, J., Bisaro, D.M., Briddon, R.W., Brown, J.K., Fauquet, C.M., Harrison, B.D., Rybicki, E.P. and Stenger, D.C. (2005). Family Geminiviridae. In Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses.

67. Sunter, G. and Bisaro, D.M. (1991). Transactivation in a geminivirus: AL2 gene product is needed for coat protein expression. *Virology*, **180**: 416-419.

68. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, **25**: 4876–4882.

69. Torres-Pacheco, I., Garzontiznado, J. A., Brown, J. K., Becerra-Flora, A. and Rivera-Bustamante, R. F. (1996). Detection and distribution of geminiviruses in Mexico and the southern United States. *Phytopathology*, **86**: 1186-1192.

70. Van Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L., Carstens, E., Estes, M. K., Lemon, S., Maniloff, J., Mayo, M. A., McGeoch, D., Pringle, C. R. and Wickner, R. B. (1999). Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press London, San Diego: Academic Press London, pp. 1044.

71. Varma, A. and Malathi, V.G. (2003). Emerging geminivirus problems: A serious threat to crop production. *Annals of Applied Biology*, **142**: 145-164.

72. Xiong, Q., Sanwei, F., Jianxiang, W. and Xueping, Z. (2007). Ageratum yellow vein China virus is a Distinct Begomovirus Species Associated with a DNAβ Molecule. *Virology*, **97**(4): 405-411.

73. Zhou, X. P., Xie, Y., Tao, X. R., Zhang, Z. K., Li, Z. H. and Fauquet, C. M. (2003). Characterization of DNA beta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNAA. *J Gen Virol*, **84**: 237–247.